Neural Networks vs Genetically Optimized Neural Networks in Time Series Prediction

Ion Railean
Monica Borda
Technical University
Cluj-Napoca, Romania

Sorin Moga
Institut TELECOM,
TELECOM Bretagne, France

Cristina Stolojescu
Politehnica University
Timisoara, Romania

Stochastic Modeling Techniques and Data Analysis International Conference,
Chania, Crete.
8 - 11 June 2010

Contents

Context of study

- Available data: WiMAX network traffic.
  - 8 weeks of complete data
  - 67 Base Stations
- The prediction model should be adaptive $\longrightarrow$ *Soft Computing Methods*.
- The Artificial Neural Networks are adapted to traffic prediction problems.

Objectives

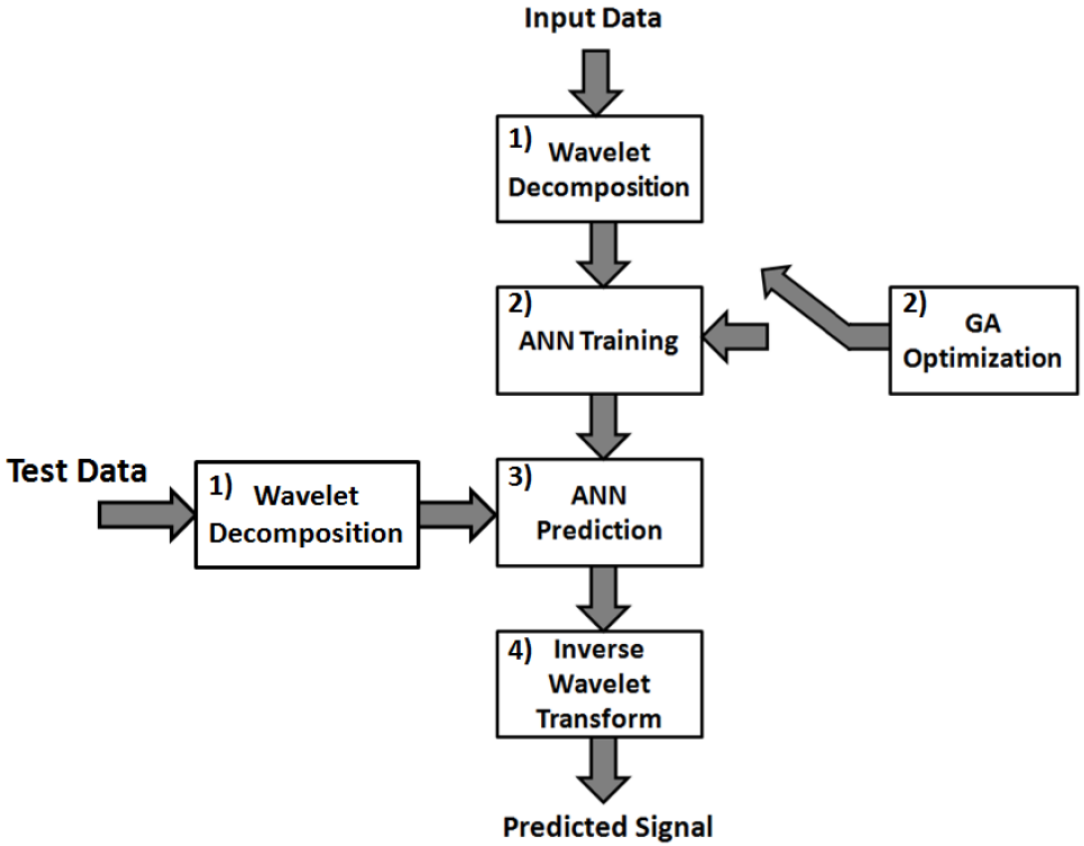
- To propose methods for finding the suitable weights of the neural networks.
- To understand the influence of the ANN optimization using GAs.
- To compare the prediction accuracy of the optimized and un-optimized ANNs.

Introduction

- The optimization of ANN using GA, applied in forecasting, has been proposed in many papers.
- Industrial processes, which do not depend on statistical and human behavior.
- In this paper : the influence of the ANN optimization using GA in a domain that implies statistical data: WiMAX network traffic.
- We combine methods from 2 research fields:
  - artificial intelligence $\longrightarrow$ the Artificial Neural Networks
  - signal processing $\longrightarrow$ the Wavelet Transform

Forecasting framework

- The simplified forecasting framework is presented in Figure 1



Wavelet Analysis

- **Multi-resolution analysis (MRA)** = signal processing technique that takes into account the signal's representation at multiple time resolutions.
- At each temporal resolution two categories of coefficients are obtained:
  - approximation coefficents.
  - details coefficients.
- We have implemented the MRA using the à trous algorithm, which corresponds to the computation of the Stationary Wavelet Transform (SWT).

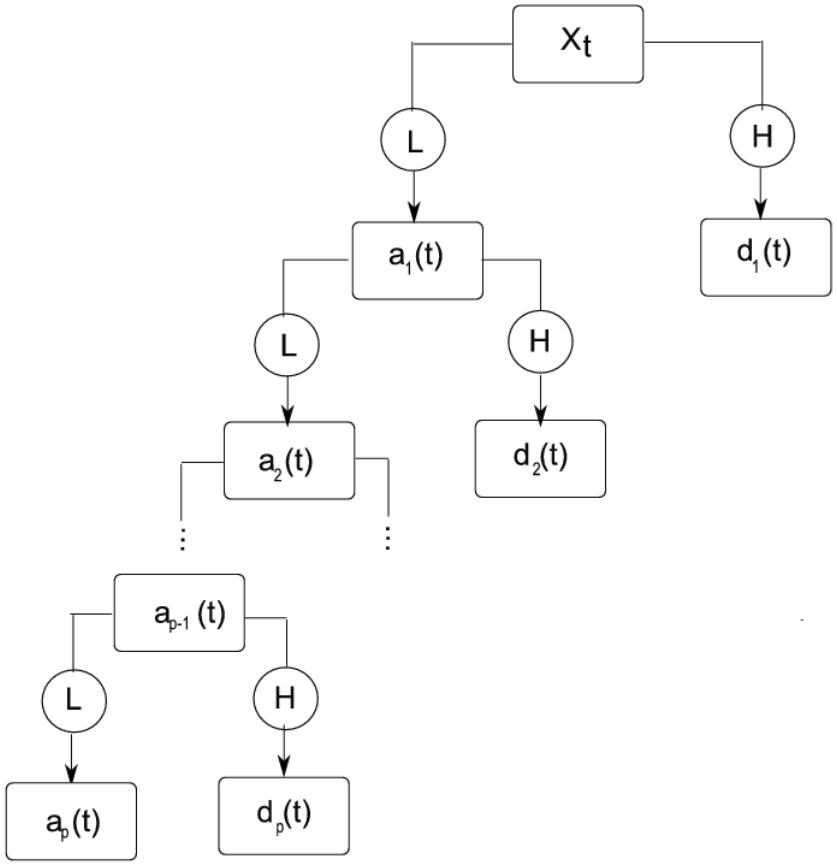
The Stationary Wavelet Transform

- two parameters:
  - the mother wavelet which generates the decomposition.
  - the number of decomposition levels.
- The SWT decomposes a signal X_t as follows:

$$X_t = a_{p,t} + \sum_{j=1}^{p} d_j, t$$

a_p, t represents the approximation at the p^{th} level of decomposition while $d_1, ...d_p$ represent the details of X_t.

SWT decomposition tree



Artificial Neural Networks

- matematical model that is composed by interconected simple elements, called neurons.
- An ANN is characterized by :
  - architecture
  - learning algorithm (training method)
  - activation function
- the most two important types of ANNs are :
  - Feed-Foreward Neural Networks
  - Racurrent Neural Networks

Artificial Neural Networks

- We used feed-forward ANN, and we discuss the setting of weights for the connections.
- Our approach consists in applying genetic algorihms to find the optimal weights between the input and the hidden layer.

Genetic Algorithms

- search technique for optimization and machine learning applications.
- a set of individual elements (the population).
- At each step:
  - the GA selects individuals randomly from the current population to be parents
  - and uses them to produce the children for the next generation.
- Over successive generations, the population ”evolves” toward an optimal solution.

Evaluation criteria

X_t - original data values, $\overline{X_t}$ - mean of X_t, F_t - predicted values
To evaluate the prediction performance between ANNs and genetically optimized ANNs, we used the following evaluation criteria:

Evaluation criteria

X_t - original data values, $\overline{X_t}$ - mean of X_t, F_t - predicted values
To evaluate the prediction performance between ANNs and genetically optimized ANNs, we used the following evaluation criteria:

- Symmetrical Mean Absolute Percentage Error (SMAPE)

$$SMAPE = \frac{2}{n} \sum_{t=1}^{n} \frac{|X_t - F_t|}{(X_t + F_t)}$$

Evaluation criteria

X_t - original data values, $\overline{X_t}$ - mean of X_t, F_t - predicted values
To evaluate the prediction performance between ANNs and genetically optimized ANNs, we used the following evaluation criteria:

- Symmetrical Mean Absolute Percentage Error (SMAPE)

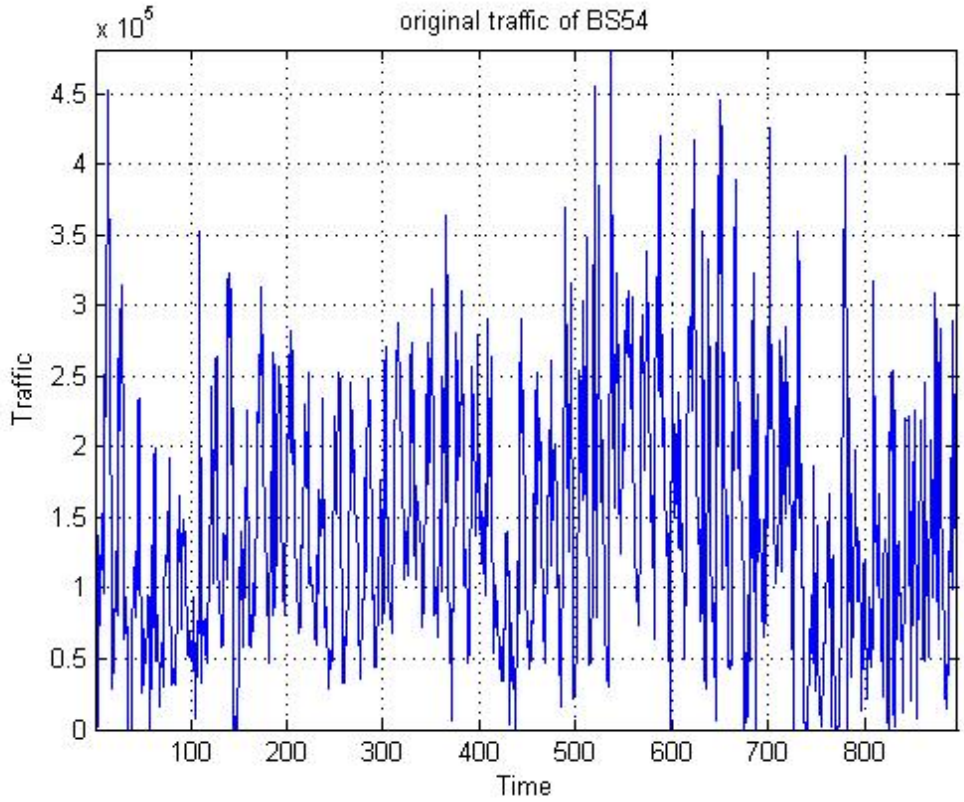
$$SMAPE = \frac{2}{n} \sum_{t=1}^{n} \frac{|X_t - F_t|}{(X_t + F_t)}$$

- R-Square (RSQ)

$$R^2 = \frac{\sum_{t=1}^{n} (F_t - \overline{F_t})^2}{\sum_{t=1}^{n} (X_t - \overline{X_t})^2}$$

Original data

- 8 weeks of complete data
- 67 Base Stations
  - up-link and down-link (bytes and packets)
  - sampled each 15 minutes (5376 samples)

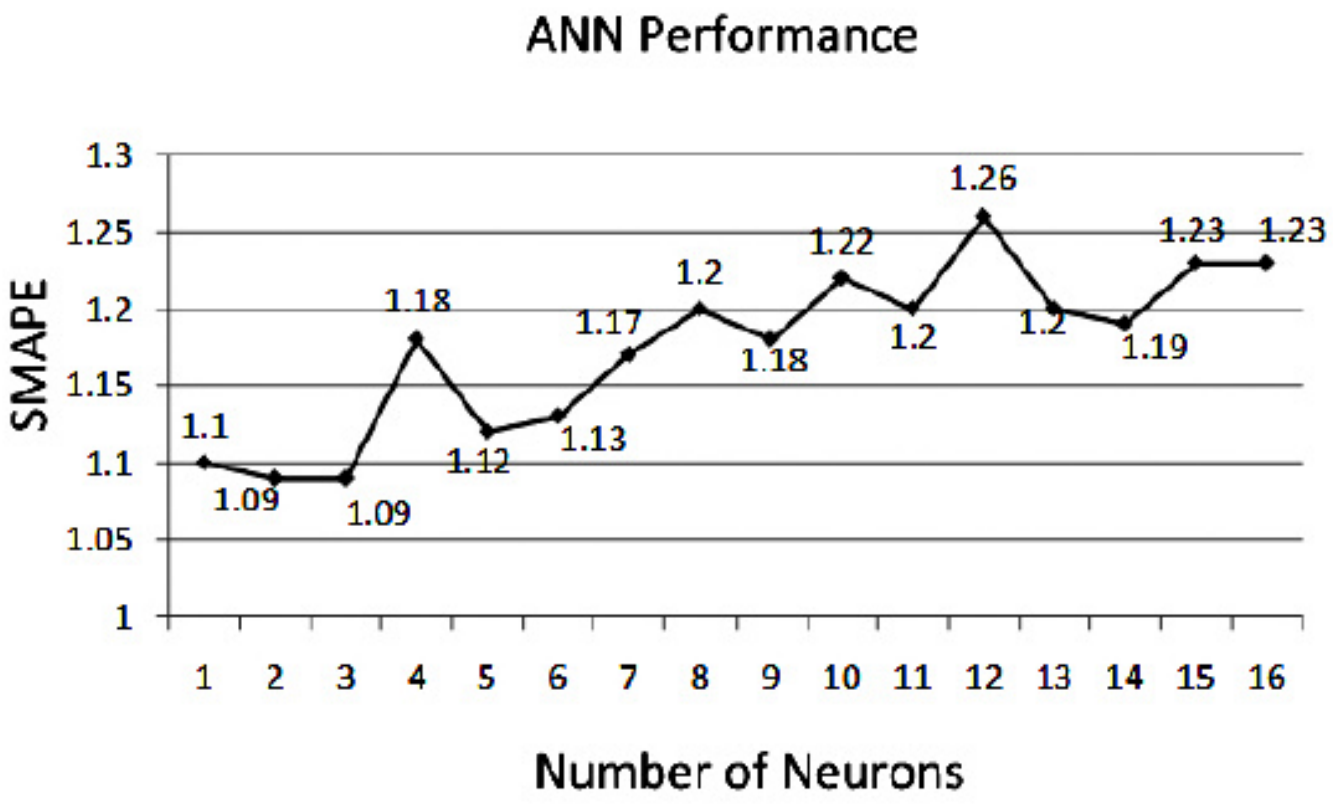


ANN parameters

- type: time-delay Neural Network(TDNN).
- output layer: one neuron with information for the targeted/predicted data.
- input layer: the time-delay of the input information was set to 4, 8, 12, and 24 hours shifting.
- number of neurons for the hidden layer: optimization.
- training algorithm: adaptive learning rate with momentum or genetic algorithm

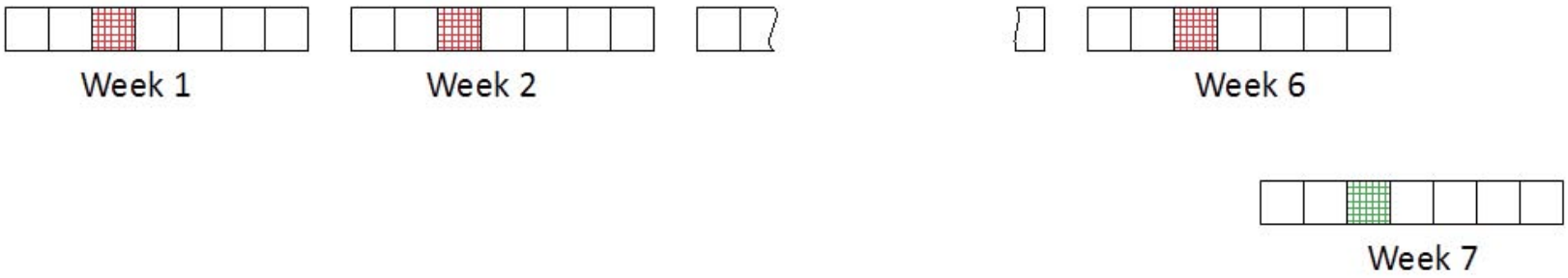
Day forecasting: ANN configuration

- Hidden layer size



Day Forecasting

- Similar Days Selection



- Take the same day of the week for processing
- Advantage: user's behavior is modeled during certain week days

Genetically Optimized ANNs approach

- two data sets at the same time.
  - (1) we used the information from the first 5 weeks while having as a target the given day from week 6.
  - (2) we used the information from weeks 2-6, while having as a target the information from week 7.
- testing phase: we applied at the inputs of the optimized ANN's the SWT corresponding to the weeks 3-7.
- The final predicted signal: inverse SWT to all forecasted sequences,
- comparison with the original signal from the 8^{th} week.

One day prediction

Time Delay	Measured Value	96 samples	32 samples	16 samples
4 hours	RSQ	1.212	1.249	1.317
	SMAPE	0.905	0.927	0.802
8 hours	RSQ	**1.159**	**1.120**	**1.205**
	SMAPE	**0.886**	**0.858**	**0.756**
12 hours	RSQ	1.330	1.219	1.287
	SMAPE	0.904	0.924	0.814
24 hours	RSQ	1.178	1.201	1.263
	SMAPE	0.910	0.946	0.780

Table: RSQ and SMAPE using ANN.

One day prediction

Time Delay	Measured Value	96 samples	32 samples	16 samples
4 hours	RSQ	1.127	1.200	1.176
	SMAPE	1.001	**0.913**	0.867
8 hours	RSQ	**1.108**	**1.129**	**1.099**
	SMAPE	**0.983**	0.924	**0.820**
12 hours	RSQ	1.211	1.155	1.168
	SMAPE	1.008	0.951	0.851
24 hours	RSQ	1.087	1.189	1.215
	SMAPE	1.067	0.950	0.872

Table: RSQ and SMAPE using Genetically Optimized ANN.

Conclusions

- The optimized ANN is able to express better the tendency and the variability of the statistical data (RSQ closer to 1).
- In case of optimized networks we have a more shifted medium value from the medium of the original signal compared to the regular ANNs.
- Regular ANN, expresses the behavior closer to the data we want to predict (SMAPE closer to 0).

Thank you for your attention!

Day Forecasting: data shifting